SDU

Deep Learning - Exercise 3

Fall 2023

To do these exercises, you will use Python with 3 and the following packages:

- NumPy. This is highly recommended package to create various data transformation / generate data.

- Pandas. With this package you can import your data into a data frame similar to how it's done in R.

- Matplotlib. This package allows you to graph your data and data transformations.

- PyTorch. This package is a versatile deep learning framework that facilitates building and training of neural networks.

- Seaborn. This package is built on top of Matplotlib, providing a high-level interface for creating informative and aesthetically appealing statistical graphics in Python.

You are not strictly forced to use these packages, but it is highly recommended. Feel free to use other packages you think are necessary.

Exercise 3

This exercise will extend on the model created during Exercise 1 and allows you to utilize your knowledge on regularization. Load the dataset by:

```python
import torch
import torch.nn as nn
import torch.optim as optim
from torchvision import datasets, transforms
import seaborn as sns
import matplotlib.pyplot as plt

# Download the MNIST dataset
transform = transforms.ToTensor()
train_dataset = datasets.MNIST(root='./data', train=True, download=True,
transform=transform)
test_dataset = datasets.MNIST(root='./data', train=False, download=True,
transform=transform)
train_loader = torch.utils.data.DataLoader(dataset=train_dataset,
batch_size=64, shuffle=True)
test_loader = torch.utils.data.DataLoader(dataset=test_dataset,
batch_size=64, shuffle=False)
```

1. Regularization:
    a. Add L1 regularization to the 2nd layer (the layer after the input layer)
    b. Add L2 regularization instead on the 2nd layer.
    c. What do you observe? (Hint: The lambda value used has a big impact on performance.)
    d. What is the purpose of adding regularization?
2. Dropout:
    a. Add a dropout layer between the first and second layer. What do you observe?
    b. What is the purpose of adding dropout?
3. Layers:
    a. Experiment with different number of layers. What do you observe?
    b. Experiment with number of neurons in each layer. What do you observe?
4. Momentum:
    a. Try to add momentum to the SGD optimizer.
    b. Test different values of momentum. With which value do you get the highest accuracy?
    c. What happens if momentum is too high?